AD-A099 144 VIRGINIA POLYTECHNIC INST AND STATE UNIV BLACKSBURG D--ETC F/G 12/1
AN ALGEBRAIC APPROACH TO SUPER-RESOLUTION ADAPTIVE ARRAY PROCES--ETC(U)
APR 81 J A CADSOW, T P BRONEZ N00014-80-C-0303

UNCLASSIFIED NL

1 OF 1
AD A
099144

END
DATE
FILMED
6-81
DTIC

AD A099144

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A099 144 / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) An Algebraic Approach to Super-Resolution Adaptive Array Processing	5. TYPE OF REPORT & PERIOD COVERED Manuscript
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) James A. Cadsow; Thomas P. Bronez	8. CONTRACT OR GRANT NUMBER(s) N00014-80-C-0303
9. PERFORMING ORGANIZATION NAME AND ADDRESS Virginia Tech, Department of Electrical Engineering, Blacksburg, VA 24061	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research (Code 436), Statistics and Probability Program, Arlington, VA 22217	12. REPORT DATE April 1981
	13. NUMBER OF PAGES 4
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

LEVEL

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Array Processing, Algebraic Characterization, Wiener Method, Pisarenko Method, Maximum Likelihood Method

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper, an algebraic characterization is made of the problem of resolving two or more closely spaced (in frequency wave number) plane waves incident on a linear array. This algebraic characterization in turn suggests a number of adaptive procedures for effecting the desired resolution. One of these procedures is herein empirically shown to provide significantly better performance when compared to other contemporary procedures used in array processing such as the Wiener filter, Pisarenko and LML algorithms. This includes both a better frequency resolving capability and a faster convergence rate.

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

411462

AN ALGEBRAIC APPROACH TO SUPER-RESOLUTION ADAPTIVE ARRAY PROCESSING

James A. Cadzow & Thomas P. Bronez

Department of Electrical Engineering
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

ABSTRACT

In this paper, an algebraic characterization is made of the problem of resolving two or more closely spaced (in frequency wave number) plane waves incident on a linear array. This algebraic characterization in turn suggests a number of adaptive procedures for affecting the desired resolution. One of these procedures is herein empirically shown to provide significantly better performance when compared to other contemporary procedures used in array processing such as the Wiener filter, Pisarenko and MLM algorithms. This includes both a better frequency resolving capability and a faster convergence rate.

I. INTRODUCTION

An important array processing problem is that of determining the directions of propagation of plane waves incident on a linear array of uniformly spaced sensors [1]. Contemporary spectral analysis has been applied to this problem and has led to the development of a variety of processing methods that are able to resolve plane waves with nearly identical directions of propagation. These methods include the Wiener Filter method [2], the Maximum Likelihood method [2], and, very recently, the Pisarenko method [3]. This paper presents an array processing approach based upon an algebraic characterization of the array processing problem. This approach is shown to encompass the methods mentioned above as well as suggesting alternate methods.

II. MODEL OF THE ARRAY DATA

Let us consider the model of multiple plane waves incident on a linear array of p sensors uniformly spaced d units apart in which the sensor measurements are contaminated by additive white noise. If there are q plane waves, it follows that at any particular instant in time, the array data $y(n)$, $0 \le n \le p-1$, has the form

$$y(n) = \eta(n) + \sum_{k=1}^{q} A_k e^{j\phi_k} e^{jn\omega_k}, \quad 0 \le n \le p-1 \tag{1}$$

where the plane wave spatial frequencies are given by

$$\omega_k = \frac{2\pi d \sin\theta_k}{\lambda}, \quad 1 \le k \le q, \tag{2}$$

and the $\{\eta(n)\}$ are uncorrelated zero mean random variables with variance σ^2, the $\{A_k\}$ are the plane waves' complex amplitudes, the $\{\phi_k\}$ are phase angles dependent on the sampling instant, the $\{\theta_k\}$ are the plane waves' directions of propagation relative to the array, and λ is the common wavelength of the plane waves. We assume that the ω_k are all different. Clearly, an estimate of the spatial frequencies ω_k directly yields an estimate of the directions of propagation $\{\theta_k\}$.

The above set of p instantaneous measurements (1) is referred to as a "snapshot". To aid the estimation of the ω_k, we utilize a number of snapshots taken sequentially in time. The array data then has the form

$$y_m(n) = \eta_m(n) + \sum_{k=1}^{q} A_k e^{j\phi_{km}} e^{jn\omega_k}, \quad 0 \le n \le p-1,\ 1 \le m \le M \tag{3}$$

where m is the snapshot index and M is the total number of snapshots used. In this model, we assume that the phase angles $\{\phi_{km}\}$ are uncorrelated random variables uniformly distributed on $[-\pi, \pi]$. This description holds due to the independence of the sinusoidal sources and from the approximate randomness of time-sampling far below the Nyquist rate.

It will be convenient to represent the given data in vector notation. The m^{th} snapshot (3) will be represented by the $p \times 1$ column vector

$$\underline{y}_m = [y_m(0) \quad y_m(1) \ \ldots \ y_m(p-1)]' . \tag{4}$$

We also define the pure complex sinusoid vector as

$$\underline{s}_j = [1 \quad e^{j\omega_j} \quad e^{j2\omega_j} \ \ldots \ e^{j(p-1)\omega_j}]' \tag{5}$$

and the noise vector associated with the m^{th} snapshot as

$$\underline{\eta}_m = [\eta_m(0) \quad \eta_m(1) \ \ldots \ \eta_m(p-1)]' . \tag{6}$$

With the above notation, we may compactly represent the snapshots (3) by the data vector

equation

$$\underline{y}_m = \underline{n}_m + \sum_{k=1}^{q} A_k e^{j\phi_{km}} \underline{s}_{\omega_k}, \quad 1 \le m \le M. \tag{7}$$

The array data $\underline{y}_m$ is random due to its dependency on the random phase angles $\{\phi_{km}\}$ and the contaminative noise $\{n_m(n)\}$. Assuming that these random variables are pairwise uncorrelated and invariant with respect to the snapshot index m, it follows that each data vector $\underline{y}_m$ can be interpreted as being a windowed realization of a wide-sense stationary random vector process. The mean value of this process is the zero vector, while its associated $p \times p$ covariance matrix is specified by

$$\underline{R} = E\{\underline{y}_m \underline{y}_m^{+}\} = \sigma^2 \underline{I}_p + \sum_{k=1}^{q} P_k\, \underline{s}_{\omega_k}\, \underline{s}_{\omega_k}^{+} \tag{8}$$

where $\underline{I}_p$ is the $p \times p$ identity matrix and $P_k = |A_k|^2$ is the power of the k^{th} plane wave. Since the random vector process is wide-sense stationary, the covariance matrix $\underline{R}$ must be positive semi-definite, Toeplitz, and Hermitian. We shall now give an algebraic approach to identifying the plane wave frequencies $\{\omega_k\}$, based upon the structure of the data $\underline{y}_m$ and the associated covariance matrix $\underline{R}$.

III. ALGEBRAIC PROCESSING APPROACH

The approach to be presented is dependent on determining a nontrivial $p \times 1$ vector $\underline{a}$ that is orthogonal to the noise-free component of each of the data vectors $\underline{y}_m$. This orthogonality is defined by the general inner product relationship

$$0 = \langle \underline{a}, \underline{y}_m - \underline{n}_m \rangle$$

$$= \sum_{k=1}^{q} A_k^{*} e^{-j\phi_{km}} \langle \underline{a}, \underline{s}_{\omega_k} \rangle, \quad 1 \le m \le M. \tag{9}$$

Since the $\{\omega_k\}$ are all different and the $\{\phi_{km}\}$ are random in nature, a little thought will convince oneself that $\underline{a}$ must be orthogonal to each of the q sinusoid vectors $\underline{s}_{\omega_k}$, $1 \le k \le q$. We next define the general z-transform $A(z)$ of the coefficient vector $\underline{a}$ by

$$A(z) = \langle \underline{a}, \underline{z}^{*} \rangle$$

where $\underline{z}' = [1 \ z^{-1} \ z^{-2} \ \ldots \ z^{1-p}]$. It is then readily shown that the orthogonality of $\underline{a}$ to each $\underline{s}_{\omega_k}$, $1 \le k \le q$, implies that $A(z)$ must have q finite zeros located on the unit circle at the points $z_k = e^{j\omega_k}$, $1 \le k \le q$. With this in mind, the required sinusoidal frequencies can be determined by examination of the zeros of $A(z)$.

Non-Idealistic Conditions

In general, if $q \ge p$ or if there is noise present, there will not exist a coefficient vector $\underline{a}$ that is orthogonal to each of the data vectors $\underline{y}_m$, $1 \le m \le M$. In either of these cases, it is intuitively desirable to select a coefficient vector which is nearly orthogonal to each of the data vectors in some well-defined manner. Once such a coefficient vector has been obtained, the plane wave frequencies are determined by examination of the zeros of the z-transform of this vector. Specifically, zeros that are close to the unit circle are considered to be indications of plane waves. Clearly, closeness is a matter of judgement; it may be conveniently evaluated by searching for nulls in the magnitude of the coefficient vector's Fourier transform as given by $A(\omega) = \langle \underline{a}, \underline{s}_\omega \rangle$.

To obtain a mathematical measure of closeness to orthogonality, it is beneficial to introduce an orthogonality error vector $\underline{e}(\underline{a})$ whose m^{th} element is the inner product of $\underline{a}$ with $\underline{y}_m$. We define the optimum $\underline{a}$ to be a vector $\underline{a}^{\circ}$ which minimizes some positive definite functional f of $\underline{e}(\underline{a})$. Hence we write

$$\underline{e}(\underline{a}) = [e(1) \quad e(2) \ \ldots \ e(M)]'$$

where

$$e(m) = \langle \underline{a}, \underline{y}_m \rangle, \tag{10}$$

and

$$f\{\underline{e}(\underline{a}^{\circ})\} = \min_{\underline{a} \in A} f\{\underline{e}(\underline{a})\} \tag{11}$$

where A is some prudently chosen set from which the solution vector $\underline{a}^{\circ}$ is to be selected.

The inner product in (10) and the functional in (11) are general at this point. We shall now choose in particular the standard vector inner product $\langle \underline{a}, \underline{y}_m \rangle = \underline{a}' \underline{y}_m^{*}$ and the normalized mean square error functional $f\{\underline{e}\} = \frac{1}{M} E\{\|\underline{e}\|^2\}$. It can be shown that

$$f\{\underline{e}(\underline{a})\} = \frac{1}{M} E\{\sum_{m=1}^{M} |\langle \underline{a}, \underline{y}_m \rangle|^2\} = \underline{a}' \underline{R}\, \underline{a} \tag{12}$$

where $\underline{R}$ is the covariance matrix (8). The functional (12) is to be minimized according to some constraint such that $\underline{a}^{\circ}$ is unique and nontrivial. Let us now consider two possible constraints.

(a) Hyperplane Constraint

The first constraint is that $\underline{a}^{\circ}$ lies on a hyperplane specified by

$$A = \{\underline{a} \in C^{p} : \underline{a}'\underline{h}^{*} + \underline{h}'\underline{a}^{*} = 2\} \tag{13}$$

where $\underline{h}$ is a nontrivial $p \times 1$ vector. The solution to (11) with this constraint can be shown to be

$$\underline{a}^{\circ} = \frac{1}{\underline{h}' \underline{R}^{-1} \underline{h}} \underline{R}^{-1} \underline{h} \tag{14}$$

and the minimum criterion's value is given by

$$f\{\underline{e}(\underline{a}^{\circ})\} = \frac{1}{\underline{h}' \underline{R}^{-1} \underline{h}}. \tag{15}$$

(b) Quadratic Constraint

The second constraint is that $\underline{a}^{\circ}$ lies on a

quadratic surface specified by

$$A = \{\underline{a} \in C^p : \underline{a}^+ \underline{W} \underline{a} = 1\} \quad (16)$$

where $\underline{W}$ is a positive definite, symmetric $p \times p$ matrix. The solution to (11) with this constraint can be shown to be

$$\underline{a}^\circ = \left(\frac{1}{\sqrt{\underline{x}_{min}^{+} \underline{W} \, \underline{x}_{min}}} \right) \underline{x}_{min} \quad (17)$$

and the minimum criterion's value is

$$f\{\underline{e}(\underline{a}^\circ)\} = \lambda_{min} \quad (18)$$

where $(\lambda_{min}, \underline{x}_{min})$ is the minimum-eigenvalue and eigenvector pair of $\underline{W}^{-1}\underline{R}$.

These two general solutions (14)-(18) encompass the three processing methods noted in the introduction: (i) For the choice $\underline{h} = [1 \ 0 \ldots 0]'$, (14) is the Wiener Filter solution [2]. As in linear prediction, this constraint implies that the first element of $\underline{a}^\circ$ is fixed and the other elements are unconstrained. (ii) For the choice $\underline{h} = \underline{s}_\omega$, (15) is the Maximum Likelihood solution [2]. This constraint implies that $A^\circ(z)$ has unity gain at $z = e^{j\omega}$ and optimally reduced gain elsewhere. (iii) For $\underline{W} = \underline{I}_p$, the quadratic surface is a hypersphere of radius one, and equation (17) is a generalization of the Pisarenko solution [3], [4]. There are several differences which distinguish this procedure from Pisarenko's. First, no ARMA model is invoked, as is done by Haykin [3]. Second, neither noise power removal nor matrix order reduction are required. Third, this solution is based upon a minimization strategy and so justifies estimates, generally even non-Toeplitz, of the covariance matrix $\underline{R}$. In the special case of a Toeplitz estimate, a power identification technique like Pisarenko's can be employed, as will be shown later. Finally, the general constraint matrix $\underline{W}$ allows greater flexibility than does the Pisarenko method.

Since the Wiener Filter solution has better resolution than the Maximum Likelihood solution [2], we shall hereafter consider only the hyperplane solution with $\underline{h} = [1 \ 0 \cdots 0]'$ and the quadratic solution with $\underline{W} = \underline{I}_p$ (hypersphere solution).

To summarize the development to this point, the algebraic approach is based on approximating an orthogonality condition between a solution vector and each of the data vectors. This approach suggests many different processing methods, depending on the choice of an inner product, an error functional, and a minimization constraint.

IV. COVARIANCE MATRIX ESTIMATE

To employ the hyperplane and hypersphere solutions given above, an estimate of the covariance matrix is required. A standard estimate is [2]

$$\hat{\underline{R}}_M = \frac{1}{M} \sum_{m=1}^{M} \underline{y}_m \, \underline{y}_m^{+} . \quad (19)$$

It is apparent that $\hat{\underline{R}}_M$ is unbiased, Hermitian, but in general not Toeplitz. Furthermore, only one lag product from each data vector is used in formulating each element of $\hat{\underline{R}}_M$. A more desirable estimate is given by the matrix $\tilde{\underline{R}}_M$ whose elements are

$$\tilde{R}_M(i,j) = c(i-j), \ 1 \le i, j \le p \quad (20)$$

where

$$c(n) = \frac{1}{M} \sum_{m=1}^{M} \frac{1}{p-n} \sum_{\ell=0}^{p-n-1} y_m(\ell+n) y_m^*(\ell) \quad , \ 0 \le n \le p-1$$

$$c(n) = c^*(-n) \ , \quad -p+1 \le n < 0 \ .$$

It is apparent that $\tilde{\underline{R}}_M$ is unbiased, Hermitian, and Toeplitz. Furthermore, it incorporates p-n lag products in formulating the covariance element c(n). Therefore the variance of $\tilde{\underline{R}}_M$ is lower than that of $\hat{\underline{R}}_M$. Thus, the estimate $\tilde{\underline{R}}_M$ is superior to the standard estimate in terms of its Toeplitz structure and lower variance.

The Toeplitz structure of $\tilde{\underline{R}}_M$ has an important implication when used with the hypersphere solution. To appreciate this, consider a general Toeplitz Hermitian matrix with a distinct minimum eigenvalue λ_{min}. An extension of Makhoul's findings [5] shows that the z-transform $X(z)$ of the eigenvector $\underline{x}$ corresponding to λ_{min} has all of its zeros located on the unit circle. Thus the *hypersphere solution will exactly indicate the* presence of p-1 plane waves if λ_{min} is distinct. Thus we have a Pisarenko-like solution and it is possible to apply a power determination technique [4], [6] to separate the q actual plane waves from the p-q-1 spurious indications (assuming $q < p$).

Given an estimate of the covariance matrix, either the hyperplane or hypersphere solutions can be employed. We now give simulation results for these different solutions.

V. SIMULATION RESULTS

To compare the performance of these processing methods, the data vectors (7) were generated by computer simulation. The simulation model corresponded to that chosen by Gabriel [2] in his comparative paper. Namely, the case of two sources incident on an array was considered. The parameter selections were $q = 2$, $p = 8$, $\sigma_1^2 = 1$, $A_1 = A_2 =$ 31.62 (30dB SNR) and 3.162 (10dB SNR), $\theta_1 = 18°$, $\theta_2 = 22°$, $d = \lambda/2$, and $M = 50$ (many snapshots) and 10 (few snapshots).

The data vectors were analyzed by four methods: the hyperplane solution with estimates $\hat{\underline{R}}_M$ and $\tilde{\underline{R}}_M$, and the hypersphere solution with $\hat{\underline{R}}_M$ and $\tilde{\underline{R}}_M$. Both the hyperplane solution with $\tilde{\underline{R}}_M$ and the hypersphere solution with $\hat{\underline{R}}_M$ showed good resolution but large spurious effects. Results for the other two methods are shown in Figure 1. In this Figure, the hyperplane solution has been evaluated via its Fourier transform and the hyper-

sphere solution has been evaluated using the power determination technique. Overlayed solutions for ten different realizations of the random data are shown to give a sense of each method's consistency.

The results show that both methods work well at the high SNR with many-snapshots. However, the hyperplane solution with $\hat{R}_M$ performs very poorly at low SNR with few snapshots, while the hypersphere solution with $\bar{R}_M$ continues to give good resolution and good suppression of spurious effects. In general, the hypersphere solution showed better performance than the hyperplane solution over a wide range of conditions.

VI. CONCLUSIONS

We have proposed an algebraic processing approach based upon approximation of a general orthogonality condition. This approach encompasses several contemporary high-resolution analysis methods. One method suggested by the algebraic approach has been shown to provide significantly better performance than other methods [2]. Further investigation of the algebraic approach is warranted in order to fully exploit its potential.

VII. REFERENCES

[1] D.E.Dudgeon, "Fundamentals of Digital Array Processing," Proc. IEEE, vol. 65, pp 898-904, 1977.

[2] W. Gabriel, "Spectral Analysis and Adaptive Array Superresolution Techniques", Proc. IEEE, vol.68, pp 654-666, June 1980.

[3] S. Haykin and J. Reilly, "Mixed Autoregressive - Moving Average Modeling of the Response of a Linear Array Antenna to Incident Plane Waves", Proc. IEEE, vol.68, pp 622-623, May 1980.

[4] V.F. Pisarenko, "The Retrieval of Harmonics from a Covariance Function", Geophys. J.R.Astr. Soc., vol.33, pp 347-366, 1973.

[5] J. Makhoul, "On the Eigenvectors of Symmetric Toeplitz Matrices", submitted for publication to IEEE Trans. Acoust., Speech, Signal Processing, February 1980.

[6] S. Haykin, ed., Nonlinear Methods of Spectral Analysis. Berlin: Springer-Verlag, 1979.

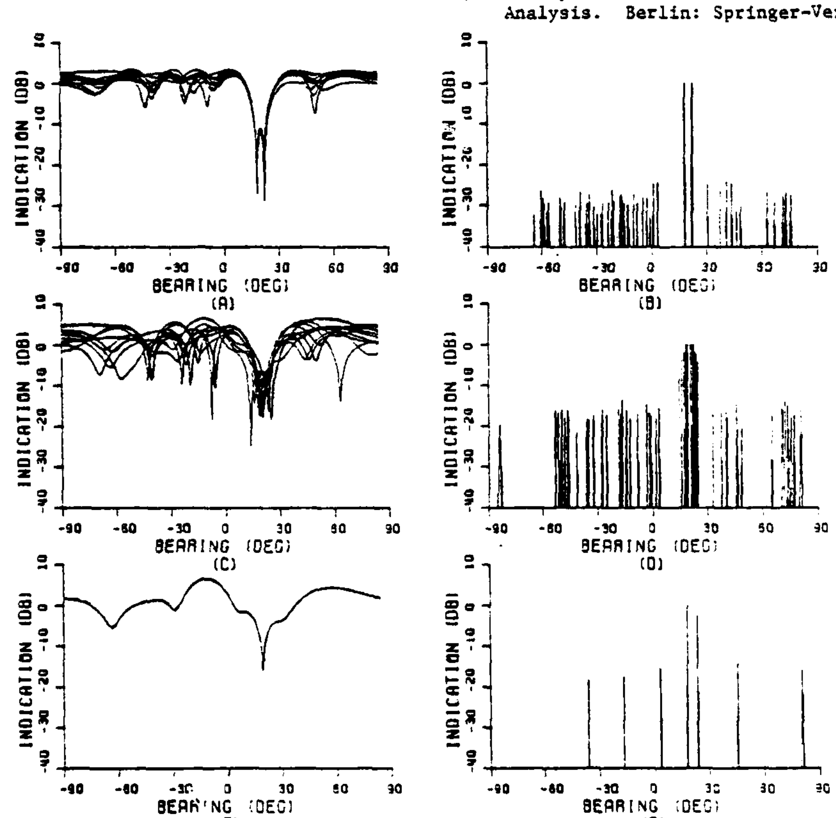

FIGURE 1. TWO-SOURCE SIMULATION WITH SOURCES AT 18 AND 22 DEGREES.
(A) HYPERPLANE SOLN. (NON-TOEP. EST.), 30DB SNR, 50 SNAPSHOTS, 10 TRIALS
(B) HYPERSPHERE SOLN. (TOEPLITZ EST.), 30DB SNR, 50 SNAPSHOTS, 10 TRIALS
(C) HYPERPLANE SOLN. (NON-TOEP. EST.), 10DB SNR, 10 SNAPSHOTS, 10 TRIALS
(D) HYPERSPHERE SOLN. (TOEPLITZ EST.), 10DB SNR, 10 SNAPSHOTS, 10 TRIALS
(E) HYPERPLANE SOLN. (NON-TOEP. EST.), 10DB SNR, 10 SNAPSHOTS, 1 TRIAL
(F) HYPERSPHERE SOLN. (TOEPLITZ EST.), 10DB SNR, 10 SNAPSHOTS, 1 TRIAL